

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/968,800DATE: 04/15/1999
TIME: 11:14:59

Input Set: H968800.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

PS

ENTERED

1 <110> APPLICANT: Hyseq, Inc.
2 <120> TITLE OF INVENTION: A NOVEL EGF MOTIF PROTEIN OBTAINED FROM
3 A cDNA LIBRARY OF FETAL LIVER-SPLEEN
4 <130> FILE REFERENCE: 20411-720
5 <140> CURRENT APPLICATION NUMBER: US/08/968,800
6 <141> CURRENT FILING DATE: 1997-11-22
7 <160> NUMBER OF SEQ ID NOS: 20
8 <170> SOFTWARE: FastSEQ for Windows Version 3.0
9 <210> SEQ ID NO 1
10 <211> LENGTH: 300
11 <212> TYPE: DNA
12 <213> ORGANISM: Homo sapiens
13 <220> FEATURE:
14 <221> NAME/KEY: CDS
15 <222> LOCATION: (1)...(300)
16 <221> NAME/KEY: misc_feature
17 <222> LOCATION: (1)...(300)
18 <223> OTHER INFORMATION: n = A,T,C or G
19 <400> SEQUENCE: 1
20 ggc tgg aga aga aac agc aag gga gtc tgt gaa gct aca tgc gaa cct 48
21 Gly Trp Arg Arg Asn Ser Lys Gly Val Cys Glu Ala Thr Cys Glu Pro
22 1 5 10 15
23 gga tgt aag ttt ggt gag tgc gtg gga cca aac aaa tgc aga tgc ttt 96
24 Gly Cys Lys Phe Gly Glu Cys Val Gly Pro Asn Lys Cys Arg Cys Phe
25 20 25 30
26 cca gga tac acc ggg aaa acc tgc agt caa gat gtg aat gag tgt gga 144
27 Pro Gly Tyr Thr Gly Lys Thr Cys Ser Gln Asp Val Asn Glu Cys Gly
28 35 40 45
29 atg aaa ccc cgg cca tgc caa cac aga tgt gtg aat aca cac gga agc 192
30 Met Lys Pro Arg Pro Cys Gln His Arg Cys Val Asn Thr His Gly Ser
31 50 55 60
32 tac aag tgc ttt tgc ctc agt ggc cac atg ctc atg cca gat gct acg 240
33 Tyr Lys Cys Phe Cys Leu Ser Gly His Met Leu Met Pro Asp Ala Thr
34 65 70 75 80
35 tgt gtg aac tcn agg aca tgt gcc atg ata aac tgt cag tat agc tgt 288
36 Cys Val Asn Xaa Arg Thr Cys Ala Met Ile Asn Cys Gln Tyr Ser Cys
37 85 90 95
38 gaa gac aca gaa 300
39 Glu Asp Thr Glu
40 100
41 <210> SEQ ID NO 2
42 <211> LENGTH: 1611
43 <212> TYPE: DNA
44 <213> ORGANISM: Homo sapiens

W-OK

PAGE: 2

RAW SEQUENCE LISTING PATENT APPLICATION US/08/968,800

DATE: 04/15/1999
TIME: 11:14:59

Input Set: H968800.RAW

45 <220> FEATURE:
46 <221> NAME/KEY: CDS
47 <222> LOCATION: (1)...(1506)
48 <221> NAME/KEY: misc_feature
49 <222> LOCATION: (1)...(1611)
50 <223> OTHER INFORMATION: n = A,T,C or G
51 <400> SEQUENCE: 2

52	ggc tgg aga aga aac agc aag gga gtc tgt gaa gct aca tgc gaa cct	48
53	Gly Trp Arg Arg Asn Ser Lys Gly Val Cys Glu Ala Thr Cys Glu Pro	
54	1 5 10 15	
55	gga tgt aag ttt ggt gag tgc gtg gga cca aac aaa tgc aga tgc ttt	96
56	Gly Cys Lys Phe Gly Glu Cys Val Gly Pro Asn Lys Cys Arg Cys Phe	
57	20 25 30	
58	cca gga tac acc ggg aaa acc tgc agt caa gat gtg aat gag tgt gga	144
59	Pro Gly Tyr Thr Gly Lys Thr Cys Ser Gln Asp Val Asn Glu Cys Gly	
60	35 40 45	
61	atg aaa ccc cgg cca tgc caa cac aga tgt gtg aat aca cac gga agc	192
62	Met Lys Pro Arg Pro Cys Gln His Arg Cys Val Asn Thr His Gly Ser	
63	50 55 60	
64	tac aag tgc ttt tgc ctc agt ggc cac atg ctc atg cca gat gct acg	240
65	Tyr Lys Cys Phe Cys Leu Ser Gly His Met Leu Met Pro Asp Ala Thr	
66	65 70 75 80	
67	tgt gtg aac tcn agg aca tgt gcc atg ata aac tgt cag tat agc tgt	288
68	Cys Val Asn Xaa Arg Thr Cys Ala Met Ile Asn Cys Gln Tyr Ser Cys	
69	85 90 95	
70	gaa gac aca gaa gaa ggg cca cag tgc ctg tgt cca tcc tca gga ctc	336
71	Glu Asp Thr Glu Glu Gly Pro Gln Cys Leu Cys Pro Ser Ser Gly Leu	
72	100 105 110	
73	cgc ctg gcc cca aat gga aga gac tgt cta gat att gat gaa tgt gcc	384
74	Arg Leu Ala Pro Asn Gly Arg Asp Cys Leu Asp Ile Asp Glu Cys Ala	
75	115 120 125	
76	tct ggt aaa gtc atc tgt ccc tac aat cga aga tgt gtg aac aca ttt	432
77	Ser Gly Lys Val Ile Cys Pro Tyr Asn Arg Arg Cys Val Asn Thr Phe	
78	130 135 140	
79	gga agc tac tac tgc aaa tgt cac att ggt ttc gaa ctg caa tat atc	480
80	Gly Ser Tyr Tyr Cys Lys Cys His Ile Gly Phe Glu Leu Gln Tyr Ile	
81	145 150 155 160	
82	agt gga cga tat gac tgt ata gat ata aat gaa tgt act atg gat agc	528
83	Ser Gly Arg Tyr Asp Cys Ile Asp Ile Asn Glu Cys Thr Met Asp Ser	
84	165 170 175	
85	cat acg tgc agc cac cat gcc aat tgc ttc aat acc caa ggg tcc ttc	576
86	His Thr Cys Ser His His Ala Asn Cys Phe Asn Thr Gln Gly Ser Phe	
87	180 185 190	
88	aag tgt aaa tgc aag cag gga tat aaa ggc aat gga ctt cgg tgt tct	624
89	Lys Cys Lys Cys Lys Gln Gly Tyr Lys Gly Asn Gly Leu Arg Cys Ser	
90	195 200 205	
91	gct atc cct gaa aat tct gtg aag gaa gtc ctc aga gca cct ggt acc	672
92	Ala Ile Pro Glu Asn Ser Val Lys Glu Val Leu Arg Ala Pro Gly Thr	
93	210 215 220	
94	atc aaa gac aga atc aag aag ttg ctt gct cac aaa aac agc atg aaa	720

W-->de

PAGE: 3

RAW SEQUENCE LISTING PATENT APPLICATION US/08/968,800

DATE: 04/15/1999
TIME: 11:14:59

Input Set: H968800.RAW

95	Ile Lys Asp Arg Ile Lys Lys Leu Leu Ala His Lys Asn Ser Met Lys	
96	225 230 235 240	
97	aag aag gca aaa att aaa aat gtt acc cca gaa ccc acc agg act cct	768
98	Lys Lys Ala Lys Ile Lys Asn Val Thr Pro Glu Pro Thr Arg Thr Pro	
99	245 250 255	
100	acc cct aag gtg aac ttg cag ccc ttc aac tat gaa gag ata gtt tcc	816
101	Thr Pro Lys Val Asn Leu Gln Pro Phe Asn Tyr Glu Glu Ile Val Ser	
102	260 265 270	
103	aga ggc ggg aac tct cat gga ggt aaa aaa ggg aat gaa gag aaa atg	864
104	Arg Gly Gly Asn Ser His Gly Gly Lys Lys Gly Asn Glu Glu Lys Met	
105	275 280 285	
106	aaa gag ggg ctt gag gat gag aaa aga gaa gag aaa gcc ctg aag aat	912
107	Lys Glu Gly Leu Glu Asp Glu Lys Arg Glu Glu Lys Ala Leu Lys Asn	
108	290 295 300	
109	gac ata gag gag cga agc ctg cga gga gat gtg ttt ttc cct aag gtg	960
110	Asp Ile Glu Glu Arg Ser Leu Arg Gly Asp Val Phe Phe Pro Lys Val	
111	305 310 315 320	
112	aat gaa gca ggt gaa ttc ggc ctg att ctg gtc caa agg aaa gcg cta	1008
113	Asn Glu Ala Gly Glu Phe Gly Leu Ile Leu Val Gln Arg Lys Ala Leu	
114	325 330 335	
115	act tcc aaa ctg gaa cat aaa gat tta aat atc tcg gtt gac tgc agc	1056
116	Thr Ser Lys Leu Glu His Lys Asp Leu Asn Ile Ser Val Asp Cys Ser	
117	340 345 350	
118	ttc aat cat ggg atc tgt gac tgg aaa cag gat aga gaa gat gat ttt	1104
119	Phe Asn His Gly Ile Cys Asp Trp Lys Gln Asp Arg Glu Asp Asp Phe	
120	355 360 365	
121	gac tgg aat cct gct gat cga gat aat gct att ggc ttc tat atg gca	1152
122	Asp Trp Asn Pro Ala Asp Arg Asp Asn Ala Ile Gly Phe Tyr Met Ala	
123	370 375 380	
124	gtt ccg gcc ttg gca ggt cac atg aaa gac att ggc cga ttg aaa ctt	1200
125	Val Pro Ala Leu Ala Gly His Met Lys Asp Ile Gly Arg Leu Lys Leu	
126	385 390 395 400	
127	ctc cta cct gac ctg caa ccc caa agc aac ttc tgt ttg ctc ttt gat	1248
128	Leu Leu Pro Asp Leu Gln Pro Gln Ser Asn Phe Cys Leu Leu Phe Asp	
129	405 410 415	
130	tac cgg ctg gcc gga gac aaa gtc ggg aaa ctt cga gtg ttt gtg aaa	1296
131	Tyr Arg Leu Ala Gly Asp Lys Val Gly Lys Leu Arg Val Phe Val Lys	
132	420 425 430	
133	aac agt aac aat gcc ctg gca tgg gag aag acc acg agt gag gat gaa	1344
134	Asn Ser Asn Asn Ala Leu Ala Trp Glu Lys Thr Thr Ser Glu Asp Glu	
135	435 440 445	
136	aag tgg aag aca ggg aaa att cag ttg tat caa gga act gat gct acc	1392
137	Lys Trp Lys Thr Gly Lys Ile Gln Leu Tyr Gln Gly Thr Asp Ala Thr	
138	450 455 460	
139	aaa agc atc att ttt gaa gca gaa cgt ggc aag ggc aaa acc ggc gaa	1440
140	Lys Ser Ile Ile Phe Glu Ala Glu Arg Gly Lys Gly Lys Thr Gly Glu	
141	465 470 475 480	
142	atc gca gtg gat ggc gtc ttg ctt gtt tca ggc tta tgt cca gat agc	1488
143	Ile Ala Val Asp Gly Val Leu Leu Val Ser Gly Leu Cys Pro Asp Ser	
144	485 490 495	

PAGE: 4

RAW SEQUENCE LISTING PATENT APPLICATION US/08/968,800

DATE: 04/15/1999
TIME: 11:14:59

Input Set: H968800.RAW

W--*OK*

```

145      ctt tta tct gtg gan nnc tgaatggtac tatctttata tttgactttg      1536
146      Leu Leu Ser Val Xaa Xaa
147              500
148      tatgtcagtt ccctgggtttt tttgatattg catcatagga cctctggcat tttaaaatta      1596
149      ctagctgaaa aattg      1611
150      <210> SEQ ID NO 3
151      <211> LENGTH: 100
152      <212> TYPE: PRT
153      <213> ORGANISM: Homo sapiens
154      <400> SEQUENCE: 3
155      Gly Trp Arg Arg Asn Ser Lys Gly Val Cys Glu Ala Thr Cys Glu Pro
156      1              5              10              15
157      Gly Cys Lys Phe Gly Glu Cys Val Gly Pro Asn Lys Cys Arg Cys Phe
158      20              25              30
159      Pro Gly Tyr Thr Gly Lys Thr Cys Ser Gln Asp Val Asn Glu Cys Gly
160      35              40              45
161      Met Lys Pro Arg Pro Cys Gln His Arg Cys Val Asn Thr His Gly Ser
162      50              55              60
163      Tyr Lys Cys Phe Cys Leu Ser Gly His Met Leu Met Pro Asp Ala Thr
164      65              70              75              80
165      Cys Val Asn Ser Arg Thr Cys Ala Met Ile Asn Cys Gln Tyr Ser Cys
166      85              90              95
167      Glu Asp Thr Glu
168              100
169      <210> SEQ ID NO 4
170      <211> LENGTH: 537
171      <212> TYPE: PRT
172      <213> ORGANISM: Homo sapiens
173      <220> FEATURE:
174      <221> NAME/KEY: VARIANT
175      <222> LOCATION: (1)...(537)
176      <223> OTHER INFORMATION: Xaa = Any Amino Acid
177      <400> SEQUENCE: 4
178      Gly Trp Arg Arg Asn Ser Lys Gly Val Cys Glu Ala Thr Cys Glu Pro
179      1              5              10              15
180      Gly Cys Lys Phe Gly Glu Cys Val Gly Pro Asn Lys Cys Arg Cys Phe
181      20              25              30
182      Pro Gly Tyr Thr Gly Lys Thr Cys Ser Gln Asp Val Asn Glu Cys Gly
183      35              40              45
184      Met Lys Pro Arg Pro Cys Gln His Arg Cys Val Asn Thr His Gly Ser
185      50              55              60
186      Tyr Lys Cys Phe Cys Leu Ser Gly His Met Leu Met Pro Asp Ala Thr
187      65              70              75              80
188      Cys Val Asn Ser Arg Thr Cys Ala Met Ile Asn Cys Gln Tyr Ser Cys
189      85              90              95
190      Glu Asp Thr Glu Glu Gly Pro Gln Cys Leu Cys Pro Ser Ser Gly Leu
191      100              105              110
192      Arg Leu Ala Pro Asn Gly Arg Asp Cys Leu Asp Ile Asp Glu Cys Ala
193      115              120              125
194      Ser Gly Lys Val Ile Cys Pro Tyr Asn Arg Arg Cys Val Asn Thr Phe

```

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/968,800

DATE: 04/15/1999

TIME: 11:14:59

Input Set: H968800.RAW

195	130	135	140	
196	Gly Ser Tyr Tyr Cys Lys Cys His Ile Gly Phe Glu Leu Gln Tyr Ile			
197	145	150	155	160
198	Ser Gly Arg Tyr Asp Cys Ile Asp Ile Asn Glu Cys Thr Met Asp Ser			
199		165	170	175
200	His Thr Cys Ser His His Ala Asn Cys Phe Asn Thr Gln Gly Ser Phe			
201		180	185	190
202	Lys Cys Lys Cys Lys Gln Gly Tyr Lys Gly Asn Gly Leu Arg Cys Ser			
203		195	200	205
204	Ala Ile Pro Glu Asn Ser Val Lys Glu Val Leu Arg Ala Pro Gly Thr			
205		210	215	220
206	Ile Lys Asp Arg Ile Lys Lys Leu Leu Ala His Lys Asn Ser Met Lys			
207	225	230	235	240
208	Lys Lys Ala Lys Ile Lys Asn Val Thr Pro Glu Pro Thr Arg Thr Pro			
209		245	250	255
210	Thr Pro Lys Val Asn Leu Gln Pro Phe Asn Tyr Glu Glu Ile Val Ser			
211		260	265	270
212	Arg Gly Gly Asn Ser His Gly Gly Lys Lys Gly Asn Glu Glu Lys Met			
213		275	280	285
214	Lys Glu Gly Leu Glu Asp Glu Lys Arg Glu Glu Lys Ala Leu Lys Asn			
215		290	295	300
216	Asp Ile Glu Glu Arg Ser Leu Arg Gly Asp Val Phe Phe Pro Lys Val			
217	305	310	315	320
218	Asn Glu Ala Gly Glu Phe Gly Leu Ile Leu Val Gln Arg Lys Ala Leu			
219		325	330	335
220	Thr Ser Lys Leu Glu His Lys Asp Leu Asn Ile Ser Val Asp Cys Ser			
221		340	345	350
222	Phe Asn His Gly Ile Cys Asp Trp Lys Gln Asp Arg Glu Asp Asp Phe			
223		355	360	365
224	Asp Trp Asn Pro Ala Asp Arg Asp Asn Ala Ile Gly Phe Tyr Met Ala			
225		370	375	380
226	Val Pro Ala Leu Ala Gly His Met Lys Asp Ile Gly Arg Leu Lys Leu			
227	385	390	395	400
228	Leu Leu Pro Asp Leu Gln Pro Gln Ser Asn Phe Cys Leu Leu Phe Asp			
229		405	410	415
230	Tyr Arg Leu Ala Gly Asp Lys Val Gly Lys Leu Arg Val Phe Val Lys			
231		420	425	430
232	Asn Ser Asn Asn Ala Leu Ala Trp Glu Lys Thr Thr Ser Glu Asp Glu			
233		435	440	445
234	Lys Trp Lys Thr Gly Lys Ile Gln Leu Tyr Gln Gly Thr Asp Ala Thr			
235		450	455	460
236	Lys Ser Ile Ile Phe Glu Ala Glu Arg Gly Lys Gly Lys Thr Gly Glu			
237	465	470	475	480
238	Ile Ala Val Asp Gly Val Leu Leu Val Ser Gly Leu Cys Pro Asp Ser			
239		485	490	495
240	Leu Leu Ser Val Asp Asp Xaa Met Val Leu Ser Leu Tyr Leu Thr Leu			
241		500	505	510
242	Tyr Val Ser Ser Leu Val Phe Leu Ile Leu His His Arg Thr Ser Gly			
243		515	520	525
244	Ile Leu Lys Leu Leu Ala Glu Lys Leu			

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Input Set: H968800.RAW

Line	? Error/Warning	Original Text
35	W "N" or "Xaa" used: Feature required	tgt gtg aac tcn agg aca tgt gcc atg ata a
67	W "N" or "Xaa" used: Feature required	tgt gtg aac tcn agg aca tgt gcc atg ata a
145	W "N" or "Xaa" used: Feature required	ctt tta tct gtg gan nnc tgaatggtac tatctt
240	W "N" or "Xaa" used: Feature required	Leu Leu Ser Val Asp Asp Xaa Met Val Leu S
255	W "N" or "Xaa" used: Feature required	Ile Asp Glu Cys Xaa Ser Asn Pro Cys Gln A
257	W "N" or "Xaa" used: Feature required	Xaa Xaa Asp Xaa Val Gly Ser Tyr Xaa Cys X
259	W "N" or "Xaa" used: Feature required	Thr Gly Lys Xaa Xaa Xaa Cys Glu Xaa Asn
270	W "N" or "Xaa" used: Feature required	Xaa Asn Glu Cys Thr Met Xaa Xaa Xaa Cys G
272	W "N" or "Xaa" used: Feature required	Val Asn Thr Xaa Gly Ser Tyr Xaa Cys Lys C
274	W "N" or "Xaa" used: Feature required	Gly Xaa Xaa Leu Xaa Cys Asp
311	W "N" or "Xaa" used: Feature required	Val Xaa Glu Cys Xaa Ser Gly Xaa Gln Xaa X
313	W "N" or "Xaa" used: Feature required	Xaa Cys Xaa Asn Thr Val Gly Ser Tyr Xaa C
315	W "N" or "Xaa" used: Feature required	Trp Xaa Pro Xaa Pro Gly Xaa Pro Asn Xaa X
330	W "N" or "Xaa" used: Feature required	Cys Val Val Gly Tyr Ile Xaa Xaa Xaa Gly G
347	W "N" or "Xaa" used: Feature required	cctttttntt tttgg
361	W "N" or "Xaa" used: Feature required	ccaaaaanaa aaagg